RAPPEL

DES PRINCIPALES LOIS DE

DISTRIBUTIONS DE

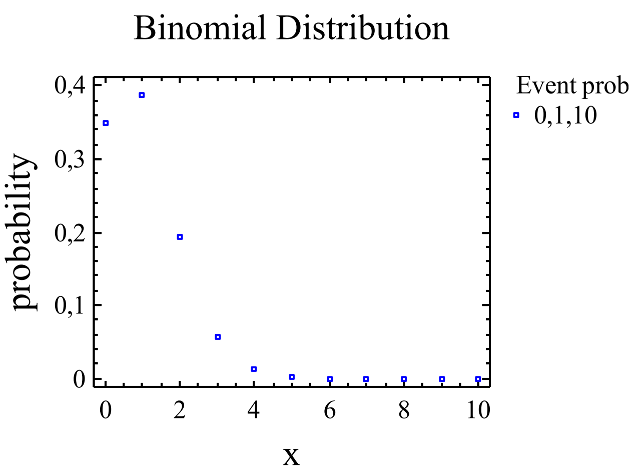
PROBABILITES

1. Loi Binomiale B(n,p)

Il s'agit en fait de la somme de n variables aléatoires indépendantes X_i qui suivent une loi $[0,1]$, de même paramètre *p*.

- $P[X = x_i] = C_{x_i}^n p^{x_i} q^{n-x_i}$ avec $q = 1 - p$
- $E(X) = np$
- $V(X) = npq$

***Somme de lois binomiales*.**
Soient deux variables aléatoires indépendantes: X_1 qui suit $B(n_1, p)$, et X_2 qui suit $B(n_2, p)$. La V.A. définie par $X = X_1 + X_2$ suit une loi $B(n_1 + n_2, p)$.



2. Loi de Poisson P(λ)

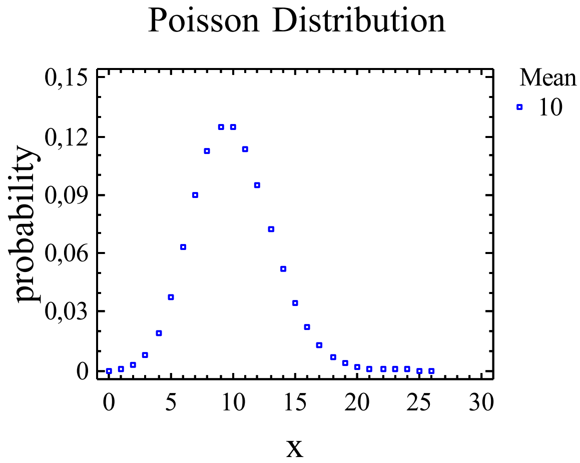
On appelle variable aléatoire de Poisson une variable aléatoire discrète X pouvant prendre des valeurs entières 0, 1, 2, …, k avec des probabilités

$$e^{-\lambda},\ \frac{\lambda}{1!} e^{-\lambda},\ \frac{\lambda}{2!} e^{-\lambda},\ \dots\ \frac{\lambda}{k!} e^{-\lambda}$$ où λ est un paramètre positif arbitraire.

- $P[X = x_i] = e^{-\lambda} \frac{\lambda^{x_i}}{x_i!}$, $k \in \mathbf{R}^+$
- $E(X) = \lambda$
- $V(X) = \lambda$

Estimation.
$\hat{\lambda} = \bar{x}$, sans biais

Somme de deux lois de Poisson.
Soient X_1 et X_2 deux variables aléatoire indépendantes suivant respectivement les lois $P(\lambda_1)$ et $P(\lambda_2)$. La loi $Z = X_1 + X_2$ suit une loi $P(\lambda_1 + \lambda_2)$.



3. Loi du Khi-deux χ^2_ν

- $f_\nu(x) = \dfrac{1}{2^{\nu/2}\Gamma\left(\dfrac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}$
- $E[X] = \nu$
- $Var[X] = 2\nu$

Somme de carré de loi normales centrées réduites.

Soient ν variables aléatoires $N(0,1)$ indépendantes $X_1, X_2, \ldots, X_\nu$. La variable aléatoire $X = \sum_{i=1}^{\nu} X_i^2$ suit une loi de probabilité χ^2_ν (khi-deux à ν degrés de liberté).

Somme de lois du Khi-deux.

Soient X et Y deux variables aléatoires indépendantes suivant une loi du χ^2 de paramètres n_1 et n_2. Alors la variable aléatoire. $Z = X + Y$ suit une loi $\chi^2_{n_1+n_2}$

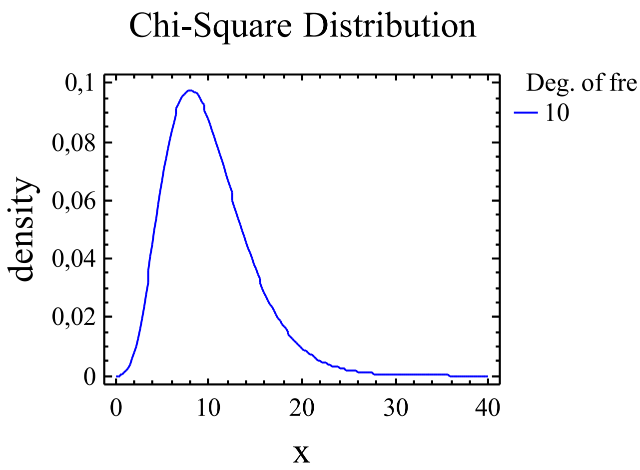
Distance du χ^2

Cette distance est utilisée pour tester si une distribution observée suit une certaine loi.

Il s'agit de la valeur $d_{\chi^2} = \sum_{\text{vecteur ou matrice}} \dfrac{(\text{distribution observée} - \text{distribution théorique})^2}{\text{distribution théorique}}$, ou plus précisément:

Pour un vecteur: $d_{\chi^2} = \sum_{i=1}^{r} \dfrac{(n_i - Np_i)^2}{Np_i}$, qui suit une loi χ^2_{r-1}. Les n_i correspondent aux quantités observées, les p_i aux probabilités théoriques (lues dans les tables), et N à la taille de l'échantillon observé.

Pour une matrice: $d_{\chi^2} = \sum_{i=1, j=1}^{r,s} \dfrac{(n_{ij} - Np_{ij})^2}{Np_{ij}}$, qui suit une loi $\chi^2_{(r-1)(s-1)}$



4. Loi Gamma $\Gamma(a, l)$

Rappel mathématique: fonction gamma d'Euler et propriétés

- $\Gamma(l) = \int_0^{+\infty} x^{l-1} \cdot e^{-x} dx$, pour $l > 0$
- $\Gamma(l+1) = l \cdot \Gamma(l)$
- $\Gamma(1) = 1$
- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$
- $\Gamma(n) = (n-1)!$, si n est entier

Généralités sur la loi gamma

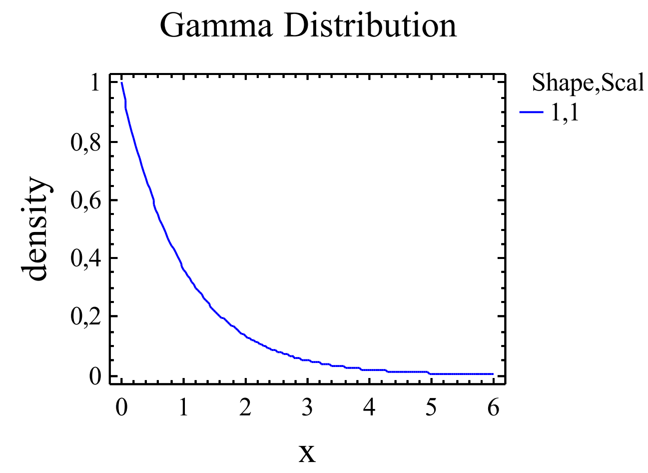
- $f(x) = \frac{a^l}{\Gamma(l)} e^{-ax} x^{l-1}$ si $x \geq 0$
- $E[X] = \frac{l}{a}$
- $V[X] = \frac{l}{a^2}$

Somme de lois gamma.

Soient X_1 et X_2, 2 variables aléatoires indépendantes suivant respectivement les lois $\Gamma(a, l_1)$ et $\Gamma(a, l_2)$. Alors la variable aléatoire $Z = X_1 + X_2$ suit une loi $\Gamma(a, l_1 + l_2)$

Rapport entre loi normale et loi gamma.

Soit X une variable aléatoire suivant la loi $N(0,1)$. Alors la loi $Y = \frac{1}{2} X^2$ suit une loi $\Gamma\left(1, \frac{1}{2}\right)$.



5. Loi Normal $N(m, \sigma^2)$

- $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}$
- $E(X) = m$
- $Var\ (X) = \sigma^2$

Moyenne.
$\hat{m} = \bar{x}$, sans biais

Variance lorsque la moyenne est connue (moyenne $m = m_0$ **).**
$\hat{\sigma}^2 = \frac{1}{n}\sum(x_i - m_0)^2$, sans biais

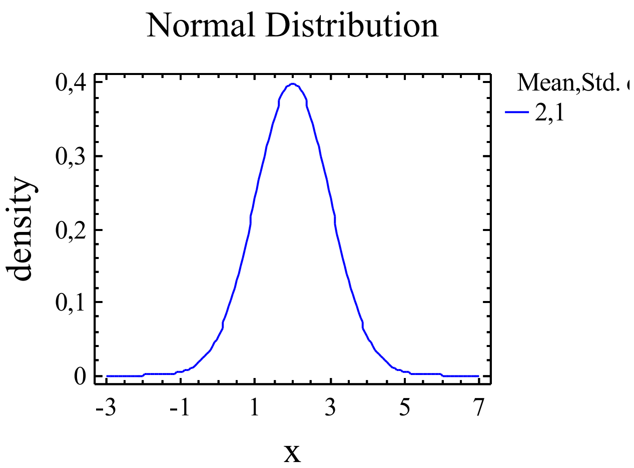
Variance lorsque la moyenne n'est pas connue.
$\hat{\sigma}^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$, biaisé. Dans ce cas, on utilise $S^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2$, sans biais. Cet estimateur prend en compte le fait qu'il faut utiliser une estimation préalable de la moyenne pour estimer la variance. Il n'y a donc plus n données disponibles (ou degrés de liberté), mais $n-1$.
<u>Remarque</u> : $S^2 = \frac{n}{n-1}\sigma^2$, σ^2 étant la valeur observée de la variance.

Centrer-réduire.
Soit une variable aléatoire X qui suit la loi normale $N(m, \sigma^2)$. La variable aléatoire $Y = \frac{X-m}{\sigma}$ suit une loi normale centrée-réduite $N(0,1)$.

La somme de p lois Normales.
$N(\mu_i, \sigma_i^2)$ deux à deux indépendantes est une loi Normale $N\left(\sum \mu_i, \sum \sigma_i^2\right)$.
Notons qu'on peut retrouver ce résultat en calculant $\sum E[X_i]$, et. $\sum V[X_i]$.

Loi de la moyenne.
Si $X_1, X_2, \ldots X_n$ deux à deux indépendants suivent une loi $N(\mu, \sigma^2)$, alors $\bar{X}$ suit une loi $N\left(\mu, \frac{\sigma^2}{n}\right)$.
Notons qu'on peut retrouver ce résultat en calculant $E[\bar{X}]$, et $V[\bar{X}]$.

Loi de la moyenne si les deux paramètres sont inconnus. Cf loi de Student.

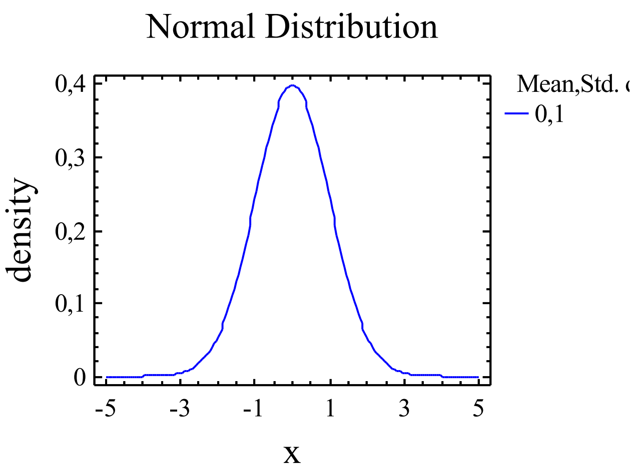


6. Loi Normale centrée réduite $N(0,1)$

- $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
- $E(X) = 0$
- $V(X) = 1$

Somme de carré de lois Normales centrées réduites (loi du Khi-deux).

Soient $X_1, X_2, \ldots X_n$ deux à deux indépendantes qui suivent chacune une loi $N(0,1)$. Alors, $Y = \sum_{i=1}^{n} X_i^2$ suit une loi χ_n^2.



7. Loi Exponentielle $\exp(\theta)$

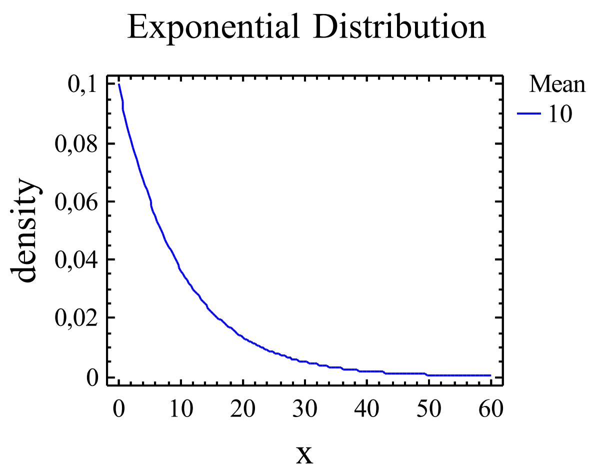
- $f_X(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$
- $E[X] = \theta$
- $Var[X] = \theta^2$

Estimation.

$\hat{\theta} = \bar{x}$

Loi Exponentielle et loi Gamma.

La loi Exponentielle $\exp(m)$ est un cas particulier de la loi Γ de paramètres $\Gamma\left(a = \frac{1}{m}, l = 1\right)$.



8. Loi de Student T_ν

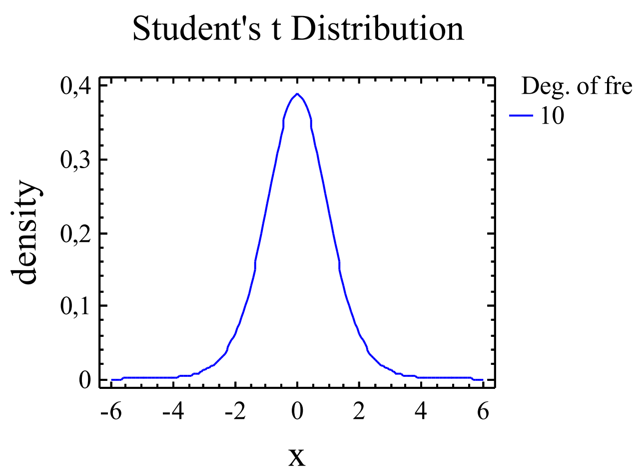
- $E[Z] = 0$
- $Var[Z] = \frac{\nu}{\nu - 2}$

Propriété.
Soient deux variables aléatoires indépendantes: X qui suit $N(0,1)$, et Y qui suit χ^2_ν. La variable aléatoire définie par $Z = \frac{X}{\sqrt{Y/\nu}}$ suit une loi T_ν (Student à ν degrés de liberté).

Loi de la moyenne d'une loi normale dont les 2 paramètres sont inconnus.
Si X suit $N(\mu, \sigma^2)$, de moyenne (théorique) $\overline{X}$, alors $\frac{\overline{X} - \mu}{S/\sqrt{n}} \rightsquigarrow T_{n-1}$ (n étant la taille de l'échantillon, et S l'estimateur de la variance). Ceci n'est valable que lorsque n est petit ($n < 30$ en pratique). Si n est plus grand, la loi de Student tend vers une loi $N(0,1)$.

Somme de variables Gaussiennes centrées.
Soient $X, X_1, X_2, \ldots, X_n$, $n+1$ variables aléatoires gaussiennes indépendantes et ***centrées***.

$T = \frac{X}{\sqrt{\frac{\sum X_i^2}{n}}}$ suit une loi T_n (Student à n degrés de liberté).



9. Loi de Bernouilli $[0,1]$

La variable aléatoire X ne peut prendre que 2 valeurs, par exemple 0 et 1 (ou échec et succès), avec les probabilités $p[X=0]=q$ et $p[X=1]=p=1-q$.

- $E[X]=p$
- $V[X]=pq$

10. Loi Discrète

Pour la valeur x d'une variable aléatoire discrète X pouvant prendre les valeurs ordonnées $x_1, x_2, \ldots, x_i, \ldots$ avec les probabilités $p_1, p_2, \ldots, p_i, \ldots$, où $\sum_1^{\infty} p_i = 1$, on notera la fonction de répartition par : $F(x_i) = \Pr[X \le x_i] = \sum_{k=1}^{i} p_k$